Speech-processing strategies designed for children

GERALD E. LOEB, MD, Sylmar, California

The design of cochlear prosthetic hardware and speech-processing strategies has been driven largely by psychophysical data from postlingually deafened adults with implants. In such subjects, success is related to the ability of the electrical stimulation to evoke the same patterns of neural activity and hence the same percepts as were produced formerly by acoustic input. Adults differ greatly in their ability to make use of information provided through electrical stimulation, particularly as temporal patterns. Recent research suggests that the manner in which information is processed in the auditory nervous system can be influenced by the type of information that is available during development of hearing. Because cochlear prostheses are used increasingly in prelingually deaf children, we must face the difficult task of designing and testing speech-processing strategies that are more appropriate for developing nervous systems whose first and only experience with sound comes from electrical stimulation. (Otolaryngol Head Neck Surg 1997;117:170-3.)

The first rule of pediatrics is that children are not small adults. Nowhere is this more true than in the nervous system. Nevertheless, all of the cochlear prostheses now being used increasingly in prelingually deaf children employ designs that attempt to recreate sensations of sound that adults would recognize from prior experience with acoustic hearing.

Clinical experience with the MultiStrategy Clarion cochlear prosthesis (Advanced Bionics Corp., Sylmar, Calif.) has revealed that adult implant users differ markedly in their preference for and ability to use various speech-processing strategies.[1] For example, most patients prefer continuous interleaved sampling (CIS) strategies with narrow biphasic pulses at fast repetition rates, but some prefer wider pulses at slower rates and others prefer compressed-analog stimulation, a completely different pattern of simultaneous, quasisinusoidal waveforms. These differences between users of the same device may arise from differences in the pathophysiology of their deafness or in the manner in which their central nervous systems originally learned to decode the complex and partially redundant cues that can be used to distinguish speech sounds.

Prelingually deafened children are more homogeneous than adults in that they have no preconceived notions about how to interpret temporospatial patterns of auditory nerve activity. When fitted with a cochlear prosthesis, their nervous systems can, indeed must, learn to extract the salient cues solely from whatever patterns are elicited by the electrical stimulation.[2] In so doing, however, they may become quite unlike adults who learned to hear with acoustic input.

PLASTICITY

In theory, children have the potential to extract useful information from activity patterns that might be confusing or annoying to adults. Animal experiments suggest that the developing nervous system tends to organize itself according to whatever sensory information is provided. For example, stimulation in a young deafened animal tends to preserve selectively those spiral ganglion cells that actually are activated[3]; more central structures rearrange their tonotopic tuning to increase the representation of the stimulated regions.[4,5] These changes might be expected to improve the ability of the brain to create perceptual distinctions from certain temporospatial details of the electrically evoked activity, perhaps at the expense of others.

Unfortunately, children cannot tell us what they do hear with a cochlear implant until they have had years of experience with a particular speech-processing strategy. The only guidance we have in designing such

---

From Advanced Bionics Corp.

Presented at the Sixth Symposium on Cochlear Implants in Children, Miami Beach, Fla., Feb. 2-3, 1996.

Reprint requests: Gerald E. Loeb, MD, Queen's University, Bio-Medical Engineering, Abramsky Hall, Kingston, Ontario K7L 3N6, Canada.

0194-5998/97/$5.00 + 0 **23/1/75102**

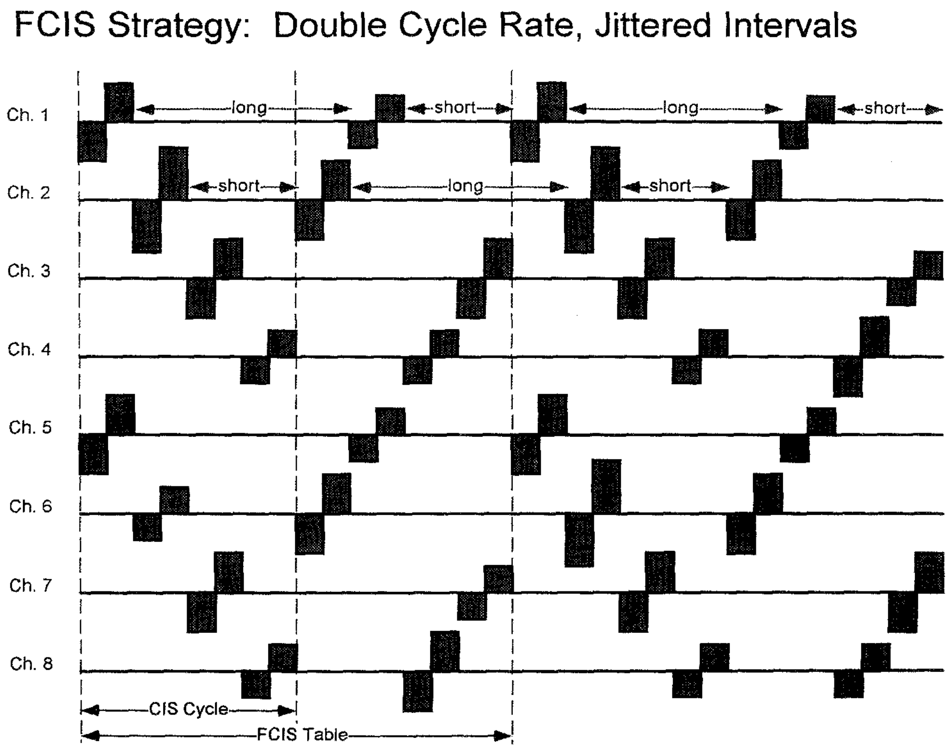

Fig. 1. Simulated output of FCIS eight-channel, double-speed strategy with jittered interpulse intervals. Strategy is encoded from table that specifies changes in output state required for each channel in each Clarion command frame (77 μsec duration, corresponding to duration of one phase of each biphasic pulse in example). To produce jittered intervals, table is two CIS cycles long, looping back on itself after execution is complete. In each CIS cycle, estimate of amplitude of sound envelope detected by each of eight band-pass filters is converted into appropriate amplitude of biphasic pulse delivered by each respective output channel. By changing sequence of output pulses from one cycle to the next in the table, two alternating interpulse intervals (long = 1000 pps; short = 2165 pps) are created on each channel. By stimulating two channels at once, rate at which new information can be encoded on all eight channels of CIS cycle is doubled from rate possible when channels are stimulated only one at a time (from 833 to 1666 Hz in example).

strategies comes from an understanding of the biophysics of intracochlear electrical stimulation, gleaned from neurophysiologic experiments in animals and psychophysical studies in adult patients with cochlear implants. Those studies suggest that the spatial information conveyed by current cochlear electrode designs varies greatly. Even the best design now available, however, is probably limited to perhaps six to 10 useful channels,[6] even when nonsimultaneous biphasic pulsing is used such as in CIS. Recent progress has focused on the quantity and fidelity of temporal information that can be provided through these channels.[7,8]

TEMPORAL ENCODING

Limitations in transmission of temporal information affect mostly the representation of the modulation envelope. In addition to providing information about rapidly changing phonemes, such as plosive consonants, the modulation envelope also provides rate-pitch cues about low frequencies that are mapped more apically than can be reached by scala tympani electrode arrays. Rate-pitch detection in adult patients with cochlear implants is poor—over 500 Hz in most patients,[9] but at least some patients can detect rate-pitch cues up to about 800 Hz. The pulsatile representation of a modulation rate requires a pulse-repetition rate that is at least twice as fast (i.e., 1600 pps). Even then, substantial distortion is introduced by aliased sampling of the acoustic signal and the tendency of auditory neurons to become phase locked to subharmonics of stimulus rates to which they are partially refractory (see below).

Increases in the stimulus-repetition rate would improve the representation of temporal information, but first a timing bottleneck must be overcome. In the CIS strategy, each channel must be stimulated in sequence with a symmetric biphasic pulse before the cycle can be repeated. For an eight-channel system operating at 1600 pps, there is less than 40 μsec available for each stimulus phase. Such narrow pulse widths require disproportionate increases in stimulus amplitude to reach adequate loudness, which increases power consumption[10]; adequate loudness may not even be attainable in some patients at the compliance voltage limit of the implant. This bottleneck can be overcome by stimulating two channels at a time, separating them spatially to minimize channel interactions. Clarion is unique among currently available cochlear prostheses in that it has eight independent stimulus generators that can be operated in any simultaneous or sequential pattern, rather than one stimulus generator that is switched rapidly between electrodes. A new family of flexible CIS (FCIS) strategies is being developed for Clarion in which the shape, sequence, and timing of pulsatile stimulation on the various channels can be set explicitly.[11]

The pulse-repetition rate that is actually required to encode temporal information remains an open question. The often-quoted Nyquist frequency used above (twice the high-frequency cutoff of the modulation frequency) is actually the minimum pulse rate required to signal simply the presence or absence of the modulation frequency. To obtain any quantitative information about the instantaneous amplitude of the modulation frequency, the stimulus-repetition rate must be higher. Even at four times the modulation frequency (e.g., 3200 pps for an 800 Hz modulator), aliasing of the samples (caused by the uncertainty of when in the waveform the amplitude measurement is taken) results in about 30% noise.

The preceding analysis assumes that the neurons respond linearly to each stimulation pulse. However, neurons have their own temporal limitations. The maximal firing rate for an individual neuron is about 600 to 800 pps. As the pulse-repetition rate approaches this maximal firing rate, the refractoriness of a neuron after responding to one stimulation pulse causes it to skip the next pulse so that it fires only at subharmonics of the pulse-repetition frequency. Further increases in stimulus pulse rate cause neurons to respond at progressively larger submultiples of the stimulus pulse rate. Because auditory neurons are fairly homogeneous, a large percentage of them can become phase-locked to the same subharmonic, effectively decreasing the rate at which envelope information is actually transmitted to the nervous system. Measurements of cochlear-evoked potentials show the phenomenon to be remarkably robust during constant amplitude stimulation at 1000 to 2000 pps.[12] Patients even report rate-pitch sensations that appear to correspond to the subharmonic apparent in the evoked potentials. However, the effect on speech information and sound quality during normally modulated speech sounds remains speculative.

One promising strategy to reduce distortion produced by neural phase locking is to introduce temporal jitter in the interpulse intervals on each channel (Fig. 1). Normally such experiments would require completely new hardware and software. However, the Multi-Strategy Clarion speech processor and clinician's fitting software are already designed to handle speech-processing strategies as different as CIS and compressed-analog stimulation. With FCIS the clinician-researcher will have complete control over the timing details of the pulsatile waveforms, including the ability to introduce timing jitter, as well as asymmetric shapes, interphase and interpulse pauses, and varying degrees of simultaneous or overlapped pulsing.

TESTING

New speech-processing strategies tend to be tested first in adults. This does allow neurophysiologists to test theories of how the spiral ganglion cells respond to electrical stimulation. Such theories provide valuable insights into biophysical processes that are likely to hold similarly for adults and children and can be used to design promising new strategies such as shown in Fig. 1. It must be remembered, however, that the percepts reported by adults are being filtered through higher centers of the auditory nervous system that have learned to interpret activity patterns in one particular way. High-level judgments and preferences of such adults may provide relatively little information about whether a child would actually find such strategies to be useful if the only auditory experience he or she had came from such stimulation. Comparing the percepts reported by adult subjects with widely differing auditory histories and subjective preferences may help to identify combinations of parameters that maximize the amount of useful information that can be presented to the auditory nervous system. Eventually, however, these "optimized" strategies must be evaluated in children who have grown up with such stimulation.

Comparisons between strategies in children necessarily take a long time because each subject must spend years learning one strategy. A prelingually deafened child who is 5 or 8 years old chronologically has the auditory nervous system of a newborn. That child will learn to hear with an implant at much the same rate that a normal newborn learns to hear. Comparisons between speech-processing strategies will also take large numbers of subjects to rule out intrinsic differences between groups of subjects. Even if the implant hardware is capable of being programmed to compare multiple strategies, a child who has learned to hear with one prosthetic strategy cannot provide a meaningful within-subject comparison of another strategy, for the same reason that postlingually deafened adult users do not provide definitive information about strategies for children.

CONCLUSIONS

A new era of pediatric cochlear prosthetics is coming. There will be strategies designed specifically for developing nervous systems rather than borrowed from fixed-design devices developed for adults. Keeping up with these strategies and their demands for signal processing, data transmission, and stimulus generation poses a major challenge for the designers and manufacturers of cochlear prostheses. Learning to select and fit such strategies wisely for their young patients poses a major challenge for clinical audiologists. Manufacturers and researchers will need to work together closely to design and evaluate systems that are practical, as well as effective.

REFERENCES

1. Kessler DK, Schindler RA. Progress with a multi-strategy cochlear implant system: the Clarion. In: Hochmair-Desoyer IJ, Hochmair ES, editors. Advances in cochlear implants. Wien: Manz; 1994. p. 354-62.
2. Loeb GE. Neural prosthetic strategies for young patients. In: Owens E, Kessler D, editors. Cochlear implants in young children. San Diego: College-Hill Press; 1989. p. 137-52.
3. Leake PA, Hrdaek GT, Rebscher SJ, Snyder RL. Chronic intracochlear electrical stimulation induces selective survival of spiral ganglion neurons in neonatally deafened cats. Hear Res 1991;54:251-71.
4. Leake PA, Snyder RL, Hradek GT, Rebscher SJ. Consequences of chronic extracochlear electrical stimulation in neonatally deafened cats. Hear Res 1995;82:65-80.
5. Snyder RL, Rebscher SJ, Cao K, Leake PA, Kelly K. Chronic intracochlear electrical stimulation in the neonatally deafened cat, I: expansion of central representation. Hear Res 1990;50:7-34.
6. Frijns JHM, de Snoo SL, ten Kate JH. Spatial selectivity in a rotationally symmetric model of the electrically stimulated cochlea. Hear Res 1996;95:33-48.
7. Shannon RV. Temporal modulation transfer functions in patients with cochlear implants. J Acoust Soc Am 1992;91:2156-64.
8. Wilson B, Finley C, Lawson D, Wolford R, Eddington D, Rabinowitz W. Better speech recognition with cochlear implants. Nature 1991;352:236-8.
9. Eddington DK, Dobelle WH, Brackmann DE, Mladejovsky MG, Parkin JL. Auditory prostheses research with multiple-channel intracochlear stimulation in man. Ann Otol Rhinol Laryngol Suppl 1978;87:5-39.
10. Loeb GE, White MW, Jenkins WM. Biophysical considerations in electrical stimulation of the auditory nervous system. Ann N Y Acad Sci 1983;405:123-36.
11. Loeb GE, Faltys M, Voelkel A. Flexible continuous interleaved sampling (FCIS): a new era for multi-strategy Clarion research. Presented at the Third International Congress on Cochlear Implants, Paris, April 27-29, 1995. p.157
12. Wilson BS, Finley CC, Zerbi M, Lawson DW. Speech processors for auditory prostheses. Final report of National Institutes of Health contract N01-DC-2-2401, 1995.